The Thesis Committee for Kathleen Marie Lyons
Certifies that this is the approved version of the following Thesis:

Exploring the Distribution of Groundwater Salamanders and Catfish with Environmental DNA

APPROVED BY
SUPERVISING COMMITTEE:

David M. Hillis, Supervisor
Justin C. Havird, Reader

Exploring the Distribution of Groundwater Salamanders and Catfish with Environmental DNA

by

Kathleen Marie Lyons

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

December 2019

Acknowledgements

This work was supported by a US Fish & Wildlife Section 6 Grant. Collaborators on the project were Dr. David Hillis, Dr. Andy Gluesenkamp, Dr. Dean Hendrickson, and Peter Sprouse. I am grateful for the participation of many private landowners, without whom sampling would not be possible. Sampling was performed with the assistance of employees of the San Antonio Zoo. The Matz lab and Cannatella lab at the University of Texas - Austin both contributed lab space so that contamination could be prevented. Probe design guidance was provided by members of the US Forest Service National Genomics Center for Wildlife and Fish Conservation.

Abstract

Exploring the Distribution of Groundwater Salamanders and Catfish with Environmental DNA

Kathleen Marie Lyons, M.A.

The University of Texas at Austin, 2019

Supervisor: David M. Hillis

Environmental-DNA (eDNA) probes were designed and tested to identify the presence of species of groundwater salamanders (genus *Eurycea*) and the Mexican blindcat (*Prietella phreatophila*) from environmental samples. Custom molecular probes were designed to identify species-specific regions of the mitochondrial *cytochrome b* gene. A new *cytochrome b* gene tree was created to ensure full coverage of all the recently revised central Texas *Eurycea* species. Successful probes that were species- or clade-specific were optimized and tested on tissue samples. Twenty-six sites across central Texas and Coahuila, Mexico, were subjected to water sampling for the purposes of eDNA analysis. These sites included both positive controls and experimental locations for both salamanders and blindcats. The presence of *P. phreatophila* was detected at a known site for the species in Val Verde County, Texas, validating both the detection method and the molecular probe. *Eurycea sp. 1* was detected at a new spring for the species close to a known sample site. An additional positive control site was Eliza Spring of the Barton Springs complex, where *E. sosorum* was detected. However, several false negative results were obtained. The development of probes for these species will aid in the discovery of new localities, and can be used to test water samples from wells and springs. Repeated sampling of localities will be needed to overcome the problem of false negative results.

Table of Contents

List of Tables

List of Figures

Background

Biological Background

Limestone karst regions feature subterranean, porous aquifers. Aquifer water can be accessed through natural features such as springs or caves, or through wells. In Central Texas and northern Mexico, human development threatens aquifer water quantity and quality (Lindgren et al., 2011). Karst aquifers provide habitat for subterranean species that are often difficult to detect, spending significant periods of time underground. As such, conservation management decisions can often be stymied by a paucity of data.

Small, aquatic salamanders of the genus *Eurycea* inhabit springs and caves over the Edwards and Trinity aquifers of central Texas. Although two *Eurycea* species — *E. sosorum* and *E. waterlooensis* — are listed as federally endangered and others are listed as threatened, designation of critical habitat can be problematic (Devitt et al., 2019). Because of the small size and difficulty of locating individuals and the amount of time spent under the surface, traditional sampling is very labor-intensive. Individual springs and caves can often be far from each other, and underground connectivity is not always well understood (Bendik et al., 2013), preventing the assignment of habitat ranges.

The blind catfish *Prietella phreatophila* (Mexican Blindcat) presents similar management problems as the Texas Eurycea. Also small and often translucent, it was known to inhabit caves, springs, and wells in Coahuila, Mexico above karst aquifers (Hendrickson et al., 2001). Recently individuals were discovered for the first time in the United States, in Amistad National Recreation Area near Del Rio, Texas (Hendrickson, et al., 2017). The Mexican Blindcat is threatened by depletion and pollution of groundwater, much like the central Texas *Eurycea*.

Methodological Background

Water, soil, and air can contain traces of biological material from living or dead organisms. Decaying tissue in the soil, pollen in the air, or the sloughed-off skin of a fish all contain what is known as environmental DNA (eDNA): genetic material either contained in whole cells or extracellular, but separate from the originating organisms. The field of eDNA analysis is relatively new, but shows promise in detecting traces of organisms that might be rare or otherwise difficult to detect through traditional sampling. Aquatic eDNA studies have traced the path of invasive fish and monitored the range of endangered crustaceans (Rees et al., 2014). Organizations like the Army Corps of Engineers and the United States Forest Service have incorporated freshwater eDNA analysis into their work.

Although freshwater eDNA sampling presents a cost-effective alternative to traditional sampling with low labor intensity, two primary concerns are contamination of samples and false negatives (Thomsen and Willerslev, 2015). All equipment must be kept as sterile as possible and the lab environment must be free of other sources of the target DNA. False negatives are common because the amount of eDNA present in any

environmental sample is extremely small, and the target DNA represents a fraction of the total eDNA. The field of eDNA detection has advanced greatly in the past decade, and analysis with quantitative PCR (qPCR) has greatly reduced the concentration threshold for detection.

Because of the miniscule concentrations of target eDNA found in aquatic samples, methods have evolved to incorporate larger and larger water samples. A significant consideration is the amount of particulate matter present in the water sample; a finer filter will capture more DNA, but will quickly clog under certain conditions (Turner et al., 2014). Because karst aquifer water is relatively clear, a finer filter can be used with less danger of clogging.

Another grand challenge of aquatic eDNA work is the possibility of determining absolute or relative concentration of the target species. Complex linear models have been used to attempt to account for persistence in the water, chemicals that can inhibit PCR, and directionality of water flow. Accounting for abundance was beyond the scope of this project, because (1) underwater aquifer flow is not well characterized in the target localities and (2) the species targeted are understood to be very rare.

Perhaps the most vital consideration after addressing contamination is primer specificity (Wilcox et al., 2013). Short molecular sequences – a forward primer, reverse primer, and fluorescent probe – bind to target regions of a gene. If the target regions do not exhibit a high level of divergence from related species, there is a significant danger of a false positive (a brown trout assay unintentionally detecting cutthroat trout). Conversely, if a target species exhibits diversity in the target region, an assay design may not account for all haplotypes and will miss some members of the group. Due to varying levels of divergence in the *cytochrome b* gene of the *Eurycea* complex, some assays designed in this study are species-specific, while others can detect a small subclade.

Methods

SAMPLE COLLECTION

Sampling occurred across the range of *Eurycea* in Texas, across the range of *Prietella phreatophila* in Coahuila, Mexico, and at the new Texas *P. phreatophila* range in Val Verde County, Texas. Sampled sites included springs, caves, and wells (Tables 1 and 2), and were divided into positive controls and experimental sites. Positive controls were sites where the target species is historically found, and some positive sites (Sotano de Amezcua) had individuals present at the time of sampling. A map of the localities is provided in Figure 1. Some sites were assayed for multiple targets, and several sites were tested both for salamanders and blindcats. At each location, between 5 and 10 L of water were pumped through 1 µm polycarbonate filter attached to a GeoTech Peristaltic Pump II. Depending on the sediment present in the water clogging the filter, additional filters were sometimes used and combined at extraction time. At springs and caves, the filter cup was placed in the water with a gloved hand, with the direction of water flowing into the cup. Marks on the filter cup were used to measure the amount of water flowing through the filter. When sampling from a well, a sterile 1L Nalgene bottle was used to collect water and pour it into the filter cup. Filters were contained in separate sterile filter cups in separate plastic bags. Pumping times (at maximum speed) ranged between 30 minutes and 1 hour. When pumping finished, the filter was transferred with clean forceps into a 2 mL tube containing 900 µL of Longmire's Buffer, which has been shown to preserve eDNA at room temperature for up to two weeks (Renshaw et al., 2015). The samples were returned as soon as possible to the University of Texas - Austin for processing. In rare cases where the peristaltic pump was unavailable, Nalgene bottles of water were brought directly to UT-Austin in a cooler and pumped through a filter using a vacuum. The complete sampling protocol can be found in Appendix A.

Table 1. *Prietella* sampling sites

Number	Date	Site	Locality	Lat	Long
Positive Controls					
1	09/2017	Sotano de Amezcua	Coahuila, Mexico	29.3199	-101.4647
2	01/2018	Catfish Parlor Cave	Val Verde County, TX	29.4714 (buffered)	-101.0325 (buffered)
3	03/2018	Cañon de Pez Blanco	Coahuila, Mexico	28.7528	-101.4303
Experimental Sites					
4	01/2018	Goodenough Spring	Val Verde County, TX	29.5361	-101.2528
5	01/2018	August Spring	Val Verde County, TX	29.5005 (buffered)	-100.9268 (buffered)
6	01/2018	San Felipe Springs #1	Val Verde County, TX	29.3728	-100.8858
7	01/2018	San Felipe Springs #3	Val Verde County, TX	29.3736	-100.8851
8	01/2018	San Felipe Springs #4	Val Verde County, TX	29.3733	-100.8834
9	01/2018	Blue Hole	Val Verde County, TX	29.5535 (buffered)	-101.0209 (buffered)
10	01/2018	ANRA-114	Val Verde County, TX	29.5118 (buffered)	-100.8895 (buffered)
11	01/2018	SNAFU Cave	Val Verde County, TX	29.5378 (buffered)	-100.9539 (buffered)
12	01/2018	Pitaya Pit	Val Verde County, TX	29.5535 (buffered)	-100.9343 (buffered)
13	03/2018	Manantial “El Chorrito”	Coahuila, Mexico	28.7675	-101.3270
14	03/2018	Cueva El Abra	Coahuila, Mexico	28.7359	-101.3796
15	03/2018	Noria Hernández Martínez, El Remolino	Coahuila, Mexico	28.7531	-101.0817
16	03/2018	Cueva Zumbadora	Coahuila, Mexico	27.1507	-101.8136

Table 2. *Eurycea* sampling sites

Number	Date	Site	Locality	Target	Lat	Long
Positive Controls						
17	05/2017	Eliza Spring	Travis County, TX	*sosorum, waterlooensis*	30.2643	-97.7702
18	09/2017	Rattlesnake Well	Hays County, TX	*rathbuni*	29.9019	-97.9214
6	01/2018	San Felipe Springs #1	Val Verde County, TX	*sp. 3*	29.3728	-100.8858
19	08/2019	Winkler Ranch Windmill Well	Blanco County, TX	*sp. 1*	30.3425	-98.2828
Experimental Sites						
20	09/2017	Geiger Well North	Hays County, TX	*pterophila, rathbuni*	29.8974	-97.9431
18	09/2017	Rattlesnake Well	Hays County, TX	*pterophila*	29.9019	-97.9214
4	01/2018	Goodenough Spring	Val Verde County, TX	*sp. 3*	29.5361	-101.2528
5	01/2018	August Spring	Val Verde County, TX	*sp. 3*	29.5005 (buffered)	-100.9268 (buffered)
7	01/2018	San Felipe Springs #3	Val Verde County, TX	*sp. 3*	29.3736	-100.8851
8	01/2018	San Felipe Springs #4	Val Verde County, TX	*sp. 3*	29.3733	-100.8834
9	01/2018	Blue Hole	Val Verde County, TX	*sp. 3*	29.5535 (buffered)	-101.0209 (buffered)
10	01/2018	ANRA-114	Val Verde County, TX	*sp. 3*	29.5118 (buffered)	-100.8895 (buffered)
11	01/2018	SNAFU Cave	Val Verde County, TX	*sp. 3*	29.5378 (buffered)	-100.9539 (buffered)
2	01/2018	Catfish Parlor Cave	Val Verde County, TX	*sp. 3*	29.4714 (buffered)	-101.0325 (buffered)
12	01/2018	Pitaya Pit	Val Verde County,	*sp. 3*	29.5535 (buffered)	-100.9343 (buffered)

			TX			
Table 2, ctd.						
3	03/2018	Cañon de Pez Blanco	Coahuila, Mexico	*sp. 3*	28.7528	-101.4303
13	03/2018	Manantial “El Chorrito”	Coahuila, Mexico	*sp. 3*	28.7675	-101.3270
14	03/2018	Cueva El Abra	Coahuila, Mexico	*sp. 3*	28.7359	-101.3796
15	03/2018	Noria Hernández Martínez, El Remolino	Coahuila, Mexico	*sp. 3*	28.7531	-101.0817
16	10/2018	Cueva Zumbadora	Coahuila, Mexico	*sp. 3*	27.1507	-101.8136
21	10/2018	Big Spring, Adams Preserve	Travis County, TX	*sp. 1*	30.3207	-98.1461
22	10/2018	Fischer Store Creek	Comal County, TX	*pterophila, rathbuni*	29.9957	-98.2194
23	11/2018	Maebius Spring 1	Hays County, TX	*sosorum*	30.2620	-98.1859
24	11/2018	Cibolo Creek Spring	Kendall County, TX	*latitans*	29.8352	-98.8433
25	11/2018	San Antonio Spring	Bexar County, TX	all *Eurycea* south of Colorado River (not *naufragia* or *tonkawae-chisholmensis*)	29.4688	-98.4673
26	04/2019	Wildomere house well	Hays County, TX	all *Eurycea* south of Colorado River (not *naufragia* or *tonkawae-chisholmensis*)	29.9111	-97.9403
27	04/2019	Wildomere stables well	Hays County, TX	all *Eurycea* south of Colorado River (not *naufragia* or *tonkawae-chisholmensis*)	29.9068	-97.9350
28	08/2019	Winkler Ranch Entrance Well	Blanco County, TX	*sp. 1*	30.3477	-98.2764
29	09/2019	West Cave	Travis County, TX	*sp. 1*	30.3409	-98.1431

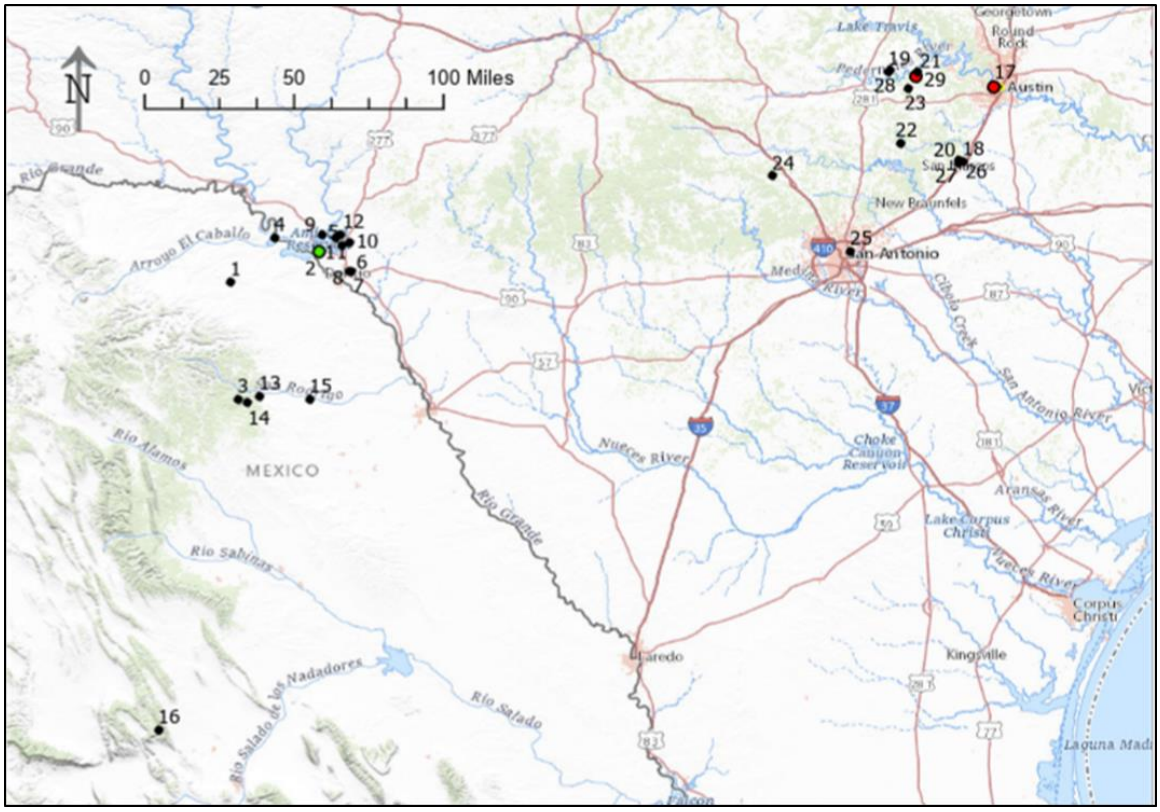


Figure 1. Map of sample localities. Numbers correspond to Tables 1 and 2. Red dots indicate positive *Eurycea* amplification and the green dot indicates positive *Prietella* amplification. Sites 2, 5, and 9 through 12 have been buffered.

SAMPLE PROCESSING

DNA extraction was performed within a week of receiving preserved filter papers. Briefly, the filter is subjected to a phenol-chloroform based extraction, including an overnight precipitation. Previous experimentation with several DNA extraction methods, filters, and commercial kits (e.g., DNEasy kit) could not ensure that all DNA was being removed from the filter. The combination of the polycarbonate filter and the phenol-chloroform extraction was chosen because the chloroform fully dissolves the polycarbonate, ensuring total DNA collection without relying on supplementary methods like bead beating (Renshaw, et al., 2015). The extraction protocol results in 200 µL of eDNA extract in LoTE buffer (dilute TE buffer), enough for several analyses if multiple probe tests are desired. The full extraction protocol is presented in Appendix A.

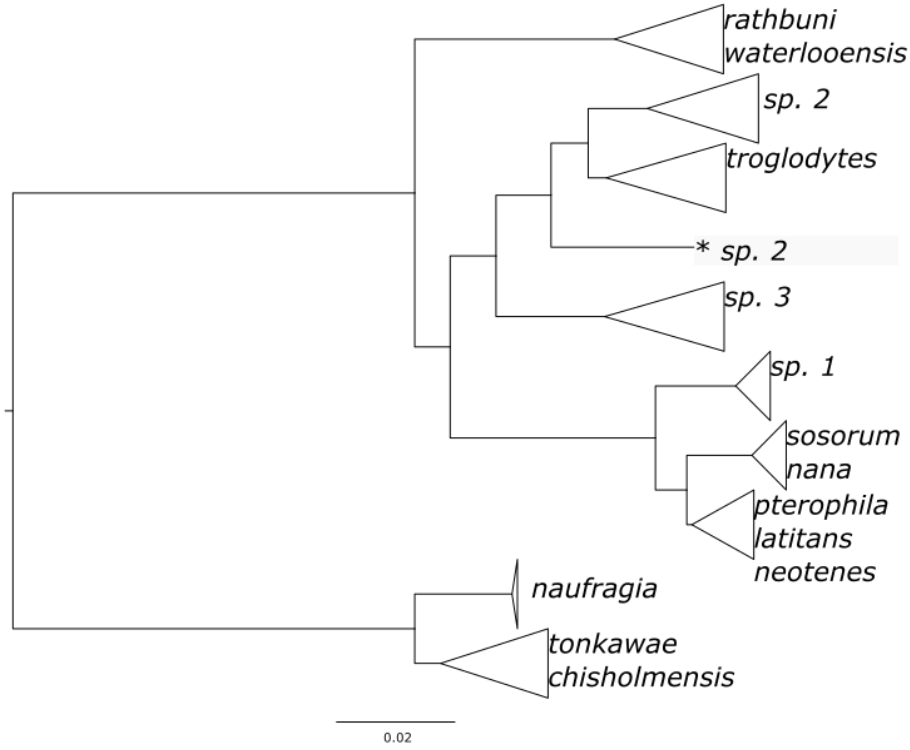
PROBE DESIGN

Primer and probe design are crucial for detecting target species in that the correct probes must amplify any member of the group without the false positives of other closely related species (Wilcox, et al., 2013). For both the *Eurycea* species and *P. phreatophila*, the target gene was the mitochondrial *cytochrome b*. Mitochondrial genes are found in eDNA at an order of magnitude greater than nuclear genes, and the *cytochrome b* (*cytb*) gene is often one of the first to be sequenced and published to repositories like GenBank. Another popular gene for eDNA assays is the *cytochrome oxidase* gene, but probes at a sufficient level of specificity could not be designed for the *Eurycea* clade. The goal of probe design was, for each target species or closely related clade, to develop a forward primer, reverse primer, and qPCR probe that would amplify a short 70-150 base pair region of the *cytb* gene. The region, or amplicon, is short because eDNA is most likely to be degraded and fragmented before collection (Turner, et al. 2014).

For each target, a collection of sequence files (in FASTA format) was collected either from GenBank or from samples sequenced in the Hillis Lab at UT-Austin. Multiple sequences from the same species or clade were included to capture the variability present. At the same time, a collection of *cytb* sequences from other amphibians (or fish, in the case of the Mexican blindcat) potentially present at the collection sites was assembled, called the nontarget library. The nontarget library included *cytb* sequences from salamanders, frogs, and toads (or fish). The target and nontarget libraries were processed using the program R (R Core Team, 2017). The package DECIPHER (Wright, 2016) contains a function called DesignPrimers. The function produced 20 forward and reverse primer pairs designed to amplify the targets, but not the nontargets. The candidate primers were analyzed using the software primer3 to eliminate primer pairs that were not within 2°C of a 60°C annealing temperature or had less than a -3 kcal/mol ΔG for any hairpins or less than a -6 kcal/mol ΔG for any dimers. This ensures that all copies of the primers will be available to bind to the target.

Once a suitable candidate primer pair was selected, TaqMan probes were designed using the program PrimerExpress. TaqMan probes are short DNA sequences with a reporter fluorophore on the 5’ end and a quencher on the 3’ end. Probes were mostly designed visually, using a trial-and-error approach. Ideal probes could only cover a region of the target where there is no known intraspecific variation, while primers are less sensitive to some mismatches. The melting temperature was 10°C higher than that of the primer pair. Probes could not have runs of more than three of the same base, and could not have a glycine on the 5’ end. Often, a sufficient probe could not be designed for a highly-ranked primer pair, and another candidate primer pair would be substituted. The full primer and probe design protocol can be found in Appendix B.

Figure 2. Mitochondrial *cytb* tree of central Texas *Eurycea*, with probes labeled. Note that a single sample of *sp. 2* (a unique haplotype found in the species) was not included in the probes design. For species that are shown together at the tips of the tree, a single probe was designed to detect the cluster of related species.



MITOCHONDRIAL GENE TREE

The *Eurycea* sites targeted by this project encompass a number of species, some of which were delimited during the course of this study (Devitt et al., 2019). To best design primers and probes that provided full coverage of the central Texas *Eurycea*, 42 additional tissues were sequenced for the *cytb* gene and a new phylogenetic tree was assembled. The primers MVZ15 and EURCB9 successfully amplified an approximately 1000 base pair region, which was Sanger sequenced at the University of Texas. In addition to the tree providing a map for probe design, some of the new sequences allowed targeting of the newest *Eurycea* species (here called *sp. 1*, *sp. 2*, and *sp. 3*). Figure 2 presents a phylogenetic tree incorporating both new and existing species using the newly sequenced tissues and older samples. With the exception of a small subclade of *Eurycea sp. 2*, all central Texas *Eurycea* are covered by a primer-probe set. The complete sequences are listed in Table 3, including the *Prietella phreatophila* assay. The placement of the *Eurycea* primer-probe sets with regards to the mitochondrial cytochrome b sequence is shown in Figure 3.

Table 3. Primers and Probes

Name	Forward Primer	Reverse Primer	Probe
sosorum-nana	AGGTGGAGTAATAGCCTTATTAGCCTCTAT	GAAATGACATGCTTCGGTGCT	CTATTCCAGCTATTCATAC
tonkawae-chisholmensis	TCTTCACGAAACCGGATCAAG	TGGGTGAAACGGGATTTGT	AACCCAACAGGACTAAA
rathbuni-waterlooensis	CACATTTGCCGCGATGTAAA	TGACGCTCCGTTAGTGTGAATATT	TACGGCTGACTTATGC
naufragia	CACGAAACCGGATCAAGCA	AACGGGATTTTGTCTGGGTTAG	CCCAACAGGACTAAAC
sp. 2	CACTAATTACACCCCCACACATTC	CGAAGGATTGCGTAGGCAAA	ACCAGAGTGATACTTCT
sp. 3	GCAACACTCACCCGGTTTTT	TGCTGACCCCTGCAATTATAAA	CTTTCCACTTTATTTTGCC
sp. 1	TCCATGAAACCGGATCAAACA	GAATGGGATTTTGTCTGGGTTAGA	CCCAACAGGAATTAA
troglodytes	ACTAACTACACCCCCACACATTCA	CGAAGGATTGCGTAGGCAAA	CCAGAGTGATACTTCTT
latitans-neotenes-pterophila	CACGAAACCGGGTCAAACA	GAATGGGATTTTGTCTGGGTTAGA	CCCAACAGGAATTAA
Prietella	CACATCTGCCGAGACGTAAA	CGCGTCCGATGTGTAGATAAA	TAGCATGTAGGTTGCGGATGAGCC

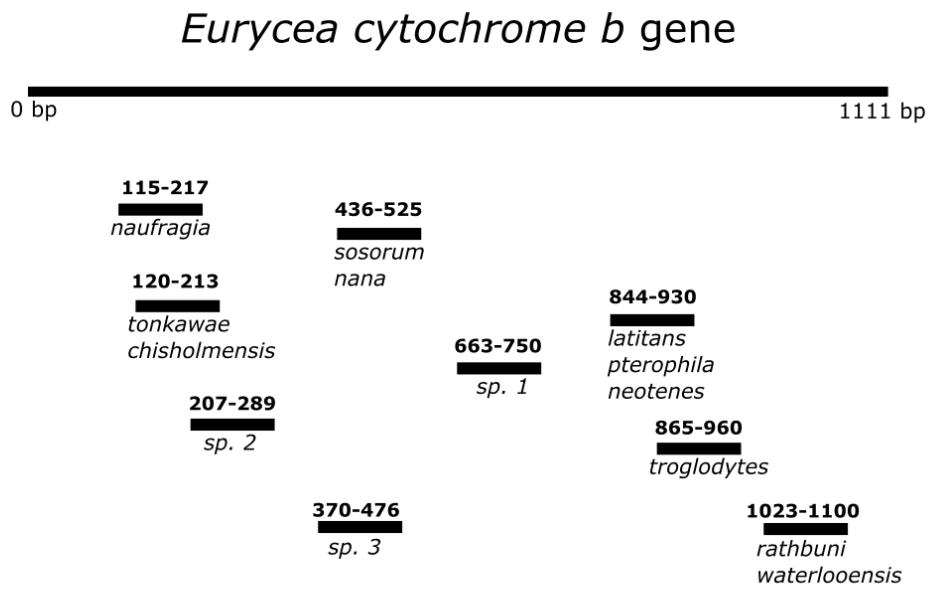


Figure 3. Schematic of *Eurycea* assays on mitochondrial *cytochrome b* gene.

PRIMER VALIDATION

Because *Eurycea* primers and probes were designed to amplify specific species or groups of closely related species, it was necessary to test whether the designed assays would cross-amplify other Eurycea targets. For example, the Eliza Spring site is home to two *Eurycea* species (*sosorum* and *waterlooensis*) covered by different assays. Representative DNA samples were selected for each *Eurycea* species, and qPCR was performed on a ViiA 7 Real-Time PCR System for 40 cycles using a SYBR Green Master Mix and a 900 nm concentration for forward and reverse primers, testing every sample against every primer set. SYBR Green is a DNA-binding dye which increases fluorescence with DNA concentration. The qPCR instrument determines a fluorescence threshold beyond which a target is said to be present, and the records the cycle number at which the threshold was reached. Instrument output is shown in Appendix C, and the results of the primer validation are shown as a heatmap in Figure 4. Each assay amplified only the target or targets, with no cross amplification. A slight amplification was noted for the *E. naufragia* sample with the *tonkawae/chisholmensis* assay, but the amplification

was late in the qPCR cycling, indicating weak annealing, and did not result in a positive detection score according to the instrument.

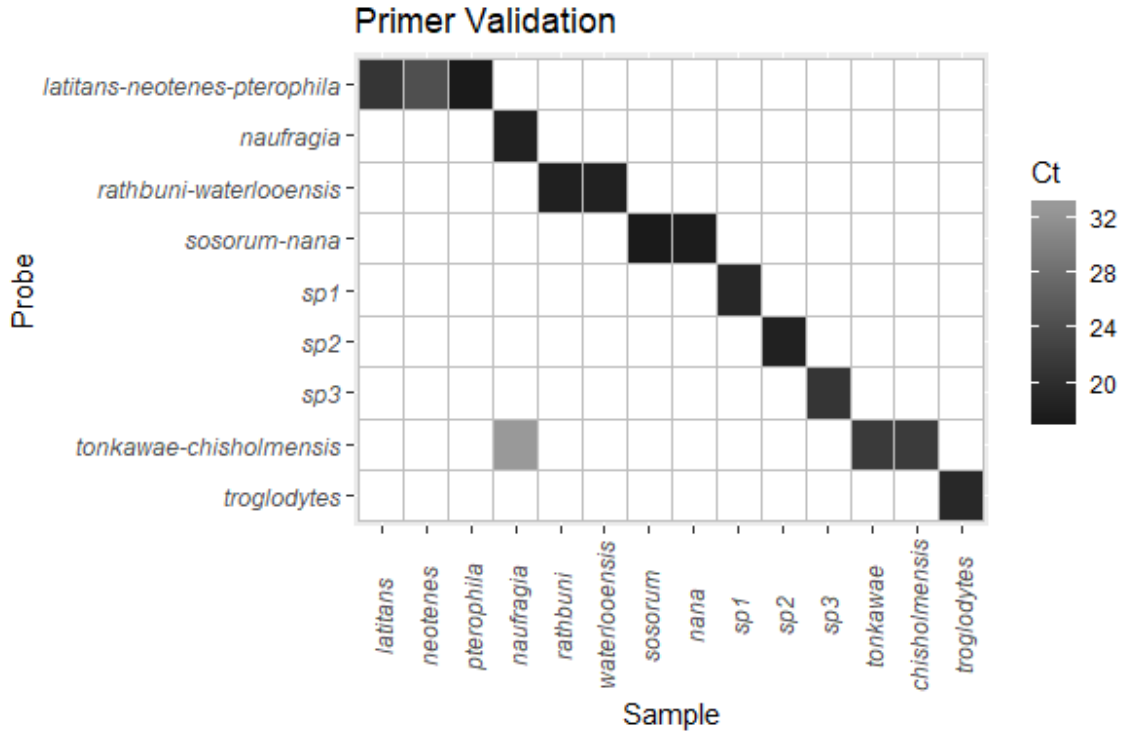


Figure 4. Primer Validation Heatmap. Ct (amplification cycle) values are plotted for Eurycea samples (x-axis) tested against all probes (y-axis). Blank squares represent a lack of amplification (true negatives). Note that the *E. naufragia* sample showed late-cycle amplification against the probe designed for its sister clade, tonkawae-chisholmensis, but the number of cycles needed for amplification exceeded our criteria for a positive detection score.

SAMPLE ANALYSIS

Samples were analyzed on a Viia-7 Real-Time PCR System using TaqMan Environmental Master Mix, a reagent set specially designed for environmental samples. Early on in the project, experimentation was performed with additives to the assay mix to attempt to neutralize common PCR inhibitors found in environmental samples. For example, EDTA was added to neutralize humic acid. However, the introduction of the TaqMan Environmental Master Mix meant that eDNA extracts did not need to be altered before analysis. Based on previous experimentation, samples were tested with assay concentrations of 900 nm for each primer and 450 nm for the probe. *Prietella* samples were run in triplicate on a 96-well plate; Eurycea samples were run on a separate plate in triplicate, but some samples were analyzed for multiple targets (see Table 1). Because the

melting temperatures of the primers and probes were similar, both the *Prietella* and *Eurycea* plates were run under the same thermocycle conditions for 60 cycles.

Results

Similar to the primer validation results, the qPCR software set a fluorescence threshold beyond which a sample is believed to be present. Triplicates were averaged and a threshold cycle (C_T) value represents the amplification cycle at which the sample crosses the fluorescence threshold value. The C_T value is presented as a table for the *Prietella* plate (Table 4) and as a heatmap for the *Eurycea* plate (Figure 5). Based on the qPCR results, a positive amplification was recorded for *Prietella phreatophila* at Catfish Parlour Cave in Val Verde County, Texas. This finding is consistent with the recent discovery of the presence of this species at this site. In the *Eurycea* plate, the Eliza Springs sample showed a positive amplification for *Eurycea sosorum*. This small spring is known for the presence of this species, and served as a positive control (although the presence of the sympatric *E. waterlooensis* was not detected). The "Big Spring" sample from Adams Preserve in Travis County, Texas also showed a positive amplification for *Eurycea sp. 1*. That species had been discovered in another spring on the property, but this spring is a new locality for the species. However, several samples are considered false negatives and are marked on Figure 5.

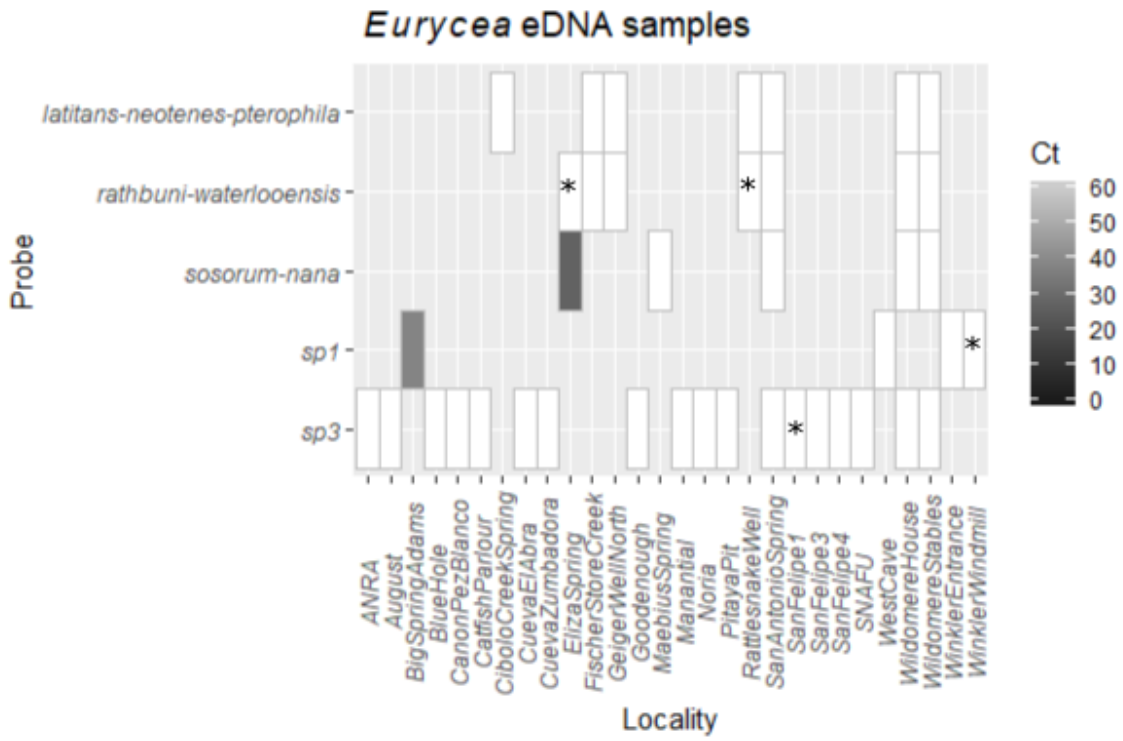


Figure 5. *Eurycea* environmental samples heatmap. White boxes represent probe-sample combinations that were tested, but failed to amplify. One positive control (Eliza Springs, *sosorum* probe) and one experimental sample (Big Spring, Adams Reserve, *sp. 1* probe) successfully amplified to give a postive signal. Asterisks mark false negatives (localities where the species is known to occur, but was not detected).

Table 4. *Prietella* qPCR results. A positive amplification was recorded for the positive control site, Catfish Parlour Cave.

Target	C_t value
Sotano de Amezcua	0
Catfish Parlour Cave	36.289
Cañon de Pez Blanco	0
Goodenough Spring	0
August Spring	0
San Felipe Springs #1	0
San Felipe Springs #3	0
San Felipe Springs #4	0
Blue Hole	0
ANRA-114	0
SNAFU Cave	0
Pitaya Pit	0
Manantial “El Chorrito”	0
Cueva El Abra	0
Noria Hernández Martínez, El Remolino	0
Cueva Zumbadora	0

Conclusions

This project established an environmental DNA sampling and analysis protocol for aquifer water (including wells). In addition, a set of molecular assays were designed encompassing species-specific or clade-specific probes for the central Texas *Eurycea* species complex and the Mexican blindcat. eDNA was successfully detected from environmental samples from both groups. The probes developed in this project will be useful to test additional sites for the presence and identification of these groups. The probe design protocol and code is customizable for any organism or gene, and the eDNA samples are available to test new targets of conservation concern, like aquifer-dwelling invertebrates.

Although several positive controls tested positive for Mexican Blindcats or expected species of *Eurycea* salamanders, other positive controls resulted in false negatives. No false positives were detected. However, it appears that false negatives are not necessarily informative – they cannot confirm the absence of a target. Because of the potential fluctuation of eDNA at a site due to biotic or abiotic factors, repeated sampling (ideally across time periods and weather conditions) could mitigate false negatives.

The false negatives may be the result of other confounding factors. For example, the water at the "Sotano de Amezcua" site was muddy and filled with particulates, and the volume of water had to be split across multiple filters, potentially diluting the DNA. The amount of both total and target eDNA may fluctuate seasonally, or even day to day (Turner et al., 2014). The field of environmental DNA is rapidly developing to include a better understanding of the ephemeral nature of aquatic eDNA to mitigate false negatives. However, repeated sampling can be difficult as aquifer water is not always flowing in caves and springs.

In conclusion, the positive amplification of eDNA at several sites demonstrates the possibility of quick, inexpensive, and low-labor detection of organisms with a low traditional detection rate. Because so many aquifer-dwellers are either of conservation concern or little-known, eDNA could be essential to determining critical habitat and establishing new conservation regimes. Stakeholders around the United States are slowly but surely incorporating eDNA into conservation assessment, and aquifers are ideal environments for this technology.

Appendices

APPENDIX A

Sampling Protocol

Supplies:
Geotech Peristaltic Pump II
plastic tubing (approximately 5 feet)
Nalgene filter funnels with original filter removed, replace with 10 µm polycarbonate filters, individually wrapped in Ziploc bags
filter adapter (sold with Nalgene funnels)
sterile forceps
gloves
2 mL tubes filled with 200 µL Longmire's solution

Protocol:
1. Remove Geotech pump, power cord, and battery from case. Use cable to hook up pump securely to battery.
2. Slide the metal piece on the pump head to open it. Insert tubing into the space created in the pump head. Slide the metal piece back to close the pump head over the tubing.
3. When the pump is set to forward, the LEFT end of the tubing is for the filter cup and the RIGHT end is for the outflow. Put the right end of the tubing in a bucket or other receptacle.
4. The white plastic filter adapter will fit over the left end of the tubing. Take a preassembled filter cup and snap it into the filter adapter. Put on gloves at this time.
5. Pour a small amount of sample water into the filter cup first. Turn on the pump and turn the speed all the way up. Continue pouring sample water into the cup.
6. Ideally collect 10L per site, with 5L being the bare minimum. It's likely that the filter will clog, making the pumping extremely slow. If this happens, just replace the filter cup with a new one and keep going.
7. When pumping is finished, the filters will need to be moved to a tube prefilled with buffer. The bottom of the filter cup separates, giving you access to the filter membrane. Take forceps (spray down with bleach if they have been previously used) and fold or roll the filter membrane. Use the forceps to insert the membrane into the tube. If multiple filters were required, you can fit 2 in the same tube.
8. Label the tube(s) with location (abbreviated), date, and number of liters pumped. In the case of multiple tubes for one location, keep them together in a labeled ziploc bag.
9. Throw out used filter cups and gloves. Spray forceps with bleach and return them to a ziploc bag.

Extraction Protocol

Supplies
hot plate or water bath set to 65°C
PCI (phenol-chloroform-isoamyl alcohol)
Sevag (chloroform-isoamyl alcohol)
5M NaCl
100% ethanol, ice cold
LoTE buffer
2 mL and 1.5 mL tubes

Protocol
1. Incubate sample tubes for 10 minutes in water bath or on hot plate.
2. Add 900 µL PCI to each tube (in fume hood).
3. Vortex tubes until filter paper is dissolved (hold tubes on their sides).
4. Centrifuge at 15,000g for 5 minutes.
5. Transfer ~700 µL aqueous later to fresh 2 mL tube.
6. Add 700 µL Sevag to tubes and vortex for 5 seconds.
7. Centrifuge at 15,000g for 5 minutes.
8. Transfer ~500 µL aqueous layer to fresh 2 mL tube.
9. Add 20 µL of 5M NaCl and 1.25 mL of ice cold 100% ethanol to tubes.
10. Precipitate at -20°C overnight.
11. Warm LoTE buffer on a hot plate or water bath.
12. Centrifuge precipitate at 15000g for 10 minutes.
13. Decant liquid from tubes.
14. Dry tubes until no liquid remains (upside down on KimWipes in fume hood).
15. Rehydrate with 200 µL LoTE buffer. May need to pipette up and down to hydrate pellet.

Generating target and nontarget libraries

For the target species (single species or group of closely related species), locate as many mitochondrial gene sequences as possible. The *cytb* or *COI* genes are widely used target genes. In-house sequencing or a resource such as GenBank should be able to provide several target sequences. It is important to use multiple sequences to capture the variation present across the target samples. For the same gene, search for nontargets: closely related species that could potentially be present in an environmental sample. For example, a probe targeting a salamander species would be designed using other local amphibians as nontargets. Using any alignment software (e.g. Muscle or MAFFT), align all the target and nontarget sequences together and generate a FASTA file. Using a text editor, manually split the sequences into separate files for targets and nontargets, and put the 2 new FASTA files in the same directory. Open R and make sure you have the DECIPHER package installed. Run the following function in the directory with the target and nontarget files.

```
getCandPrimers <- function(target, nontarget, prefix){

  #Arguments:
  # target: character       filename  of target spp fas file
  # nontarget: character    filename of nontarget spp fas file
  # prefix: character       prefix for output filenames


  require(DECIPHER)

  db <- dbConnect(SQLite(), ":memory:")

  Seqs2DB(target, "FASTA", db, "target") # Put sequences in database

  Seqs2DB(nontarget, "FASTA", db, "nontarget") # Put sequences in
database

  tiles <- TileSeqs(db, identifier="", minLength=29, maxLength=30,
verbose=T)
```

```
  write.csv(tiles, file=paste(prefix,"assay_tiles.csv", sep="_")) #Save
out into directory

  #Design the primers
  primers <- DesignPrimers(
    tiles,
    identifier="target", #targeting target
    minLength=15, #primers at least 15 bp
    maxLength=30, #primers no more than 30 bp
    maxPermutations=1, #only one permutation per primer
    minCoverage=0.9, #must have at least 90% sequence coverage
    minGroupCoverage=0.9, #must target at least 90% of targets
    annealingTemp=66, #66 C annealing Tm because program often
underestimates
    P=4e-07, #Primer concentration
    monovalent=0.07,#NULL ion concentration
    divalent=0.003, #NULL ion concentration
    dNTPs=8e-04, #NULL dNTP concentration
    worstScore=-Inf, #Do not remove poor primers from consideration
    numPrimerSets=20, #Return 20 top primer sets
    minProductSize=70, #Amplicon is at least 70 bp
    maxProductSize=150, #Amplicon is no more than 150 bp
    maxSearchSize=500, #Check for mis-priming 500 bp up and downstream
of target site
    batchSize=1000, #Pass up to 1000 primer sets to OlioArrayAux for
consideration
    maxDistance=0.4, #NULL, not full sure I understand this rule
    primerDimer=1e-07, #Maximum primer dimer efficiency
    ragged5Prime=TRUE, #Not applicable for single permutation per
primer
    taqEfficiency=TRUE, #Using Taq polymerase which has sensitivity to
3' mismatches on elongation
    verbose=TRUE) #Give status update


  #Save out these primers to this file
  #The result is a csv file (can be opened in Excel) with primer sets
ordered by specificity (most specific primer sets first)
  write.csv(primers, file = paste(prefix,"candidate_primers.csv",
sep="_"))

  return(primers)
}
```

A CSV file will appear in your directory that can be opened in Excel. It contains 20 candidate primer pairs and some details about them. The next step is to evaluate them for melt temperature, GC content, hairpins, homodimers, and heterodimers. This can be done manually by entering each sequence into the OligoAnalyzer website, but this can

take a long time. The following Python script can take that CSV file, evaluate the thermodynamic features of the primer pairs, and export a new CSV file with those features added. The Python script requires the primer3-py package to be installed.

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Thu Jun  6 17:20:20 2019

@author: klyons19
"""
#use F9 to execute lines in
import sys
import primer3
import pandas as pd

# import .csv generated from R DECIPHER, analyze primer pairs
filename = sys.argv[1]
stem = filename.split('.')[0]
data = pd.read_csv(filename)
#data = pd.read_csv('CHISNAUF_cytb_Primers.csv')
subset = data[['forward_primer','reverse_primer']].copy() #avoids
SettingWithCopyWarning
FPrimers = data['forward_primer']
RPrimers = data['reverse_primer']

subset['FTemp'] = FPrimers.apply(primer3.calcTm)


for index, row in subset.iterrows():
    subset.loc[index,'FHairpin'] = primer3.calcHairpin(row[0]).dg/1000
    subset.loc[index,'FHomodimer'] =
primer3.calcHomodimer(row[0]).dg/1000
    #print(primer3.calcHeterodimer(row[0],row[1]).dg/1000)

subset['RTemp'] = RPrimers.apply(primer3.calcTm)

for index, row in subset.iterrows():
    subset.loc[index, 'RHairpin'] = primer3.calcHairpin(row[1]).dg/1000
    subset.loc[index, 'RHomodimer'] =
primer3.calcHomodimer(row[1]).dg/1000

for index, row in subset.iterrows():
    subset.loc[index, 'FRHeterodimer'] =
primer3.calcHeterodimer(row[0],
              row[1]).dg/1000

#write csv
outfilename = stem + '_Analyzed.csv'
export_csv = subset.to_csv(outfilename)
```

Select a few candidate pairs based on the information in the new CSV. Primer pairs should be within 2°C of a 60°C annealing temperature, have between 20% and 80% GC content, have greater than a -3 kcal/mol ΔG for any hairpins, and have greater than a -6 kcal/mol ΔG for any dimers (homodimers or heterodimers). Once a short list of primer pairs has been established, probes will be designed manually. Software like primer3 or PrimerExpress or the free OligoAnalyzer website can be used to determine the properties of candidate probes.

Probes should have a melt temperature of 8-10°C greater than the primers, so about 70°C. The length can be adjusted to meet this melt temperature, but should not be shorter than 13 bp. The GC content should be in between 20% and 80%, and ideally should have more Cs than Gs. There should not be runs of more than three of the same base, and there should not be a G on the 5' end. It is most important that there are no mismatches between the target sequences and the probes, where a few mismatches with the primers can be tolerated. Probes can be ordered from Life Technologies with a 6FAM fluorophore and an MGNBFQ quencher.

The *Eurycea* and *Prietella* assays designed in this project work best with a 900 nm concentration for each primer, and a 450 nm concentration for each probe, but it may be necessary to evaluate the performance of the assay under different concentrations. Serial dilutions of the primers (for example: 100 nm, 300 nm, 600 nm, 900 nm) will reveal an optimal concentration for specificity and sensitivity.

APPENDIX C. QPCR DATA

Primer Validation Assay

Probe abbreviations:
LNP: *latitans-neotenes-pterophila*
N: *naufragia*
RW: *rathbuni-waterlooensis*
SN: *sosorum-nana*
S1: *sp. 1*
S2: *sp. 2*
S3: *sp. 3*
TC: *tonkawae-chisholmensis*
T: *troglodytes*

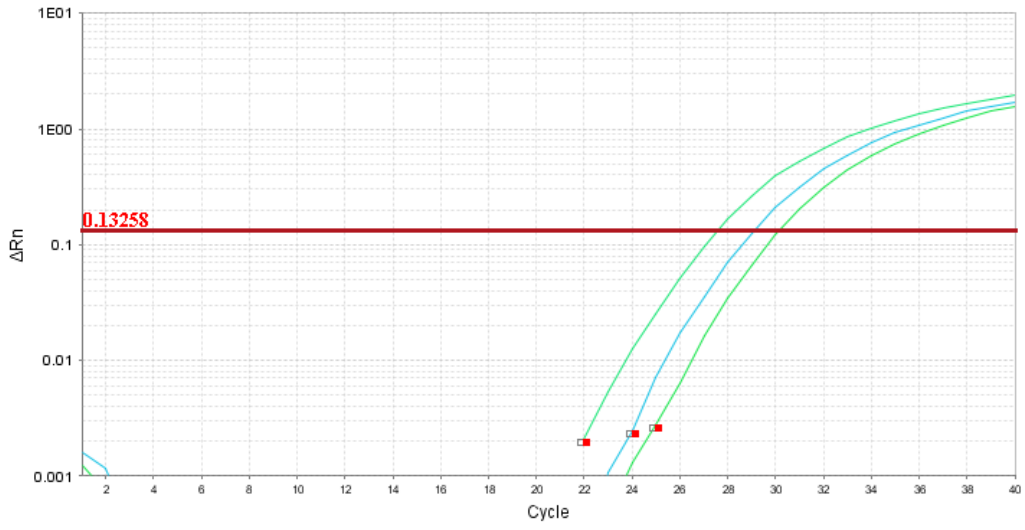


Figure 6. LNP probe with TNHC-51174 (*latitans*), TNHC-60313 (*neotenes*), TNHC-60316 (*pterophila*)

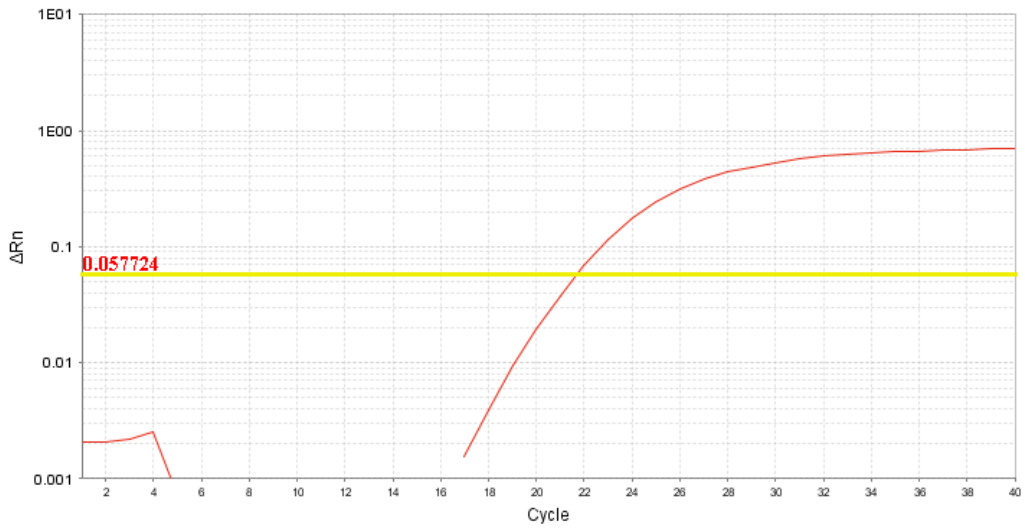


Figure 7. N probe with AGG-2020 (*naufragia*)

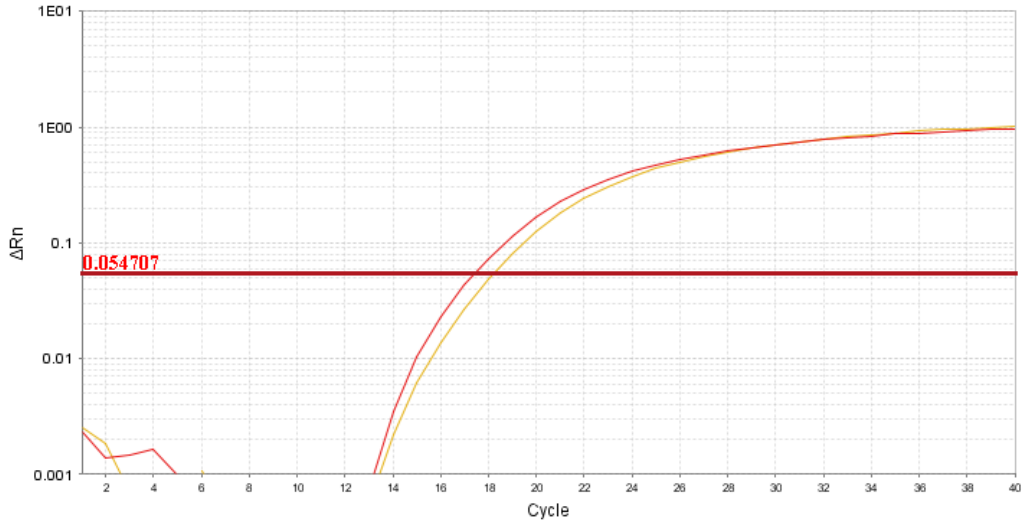


Figure 8. RW probe with AGG-1981 (*rathbuni*), AGG-1994 (*waterlooensis*)

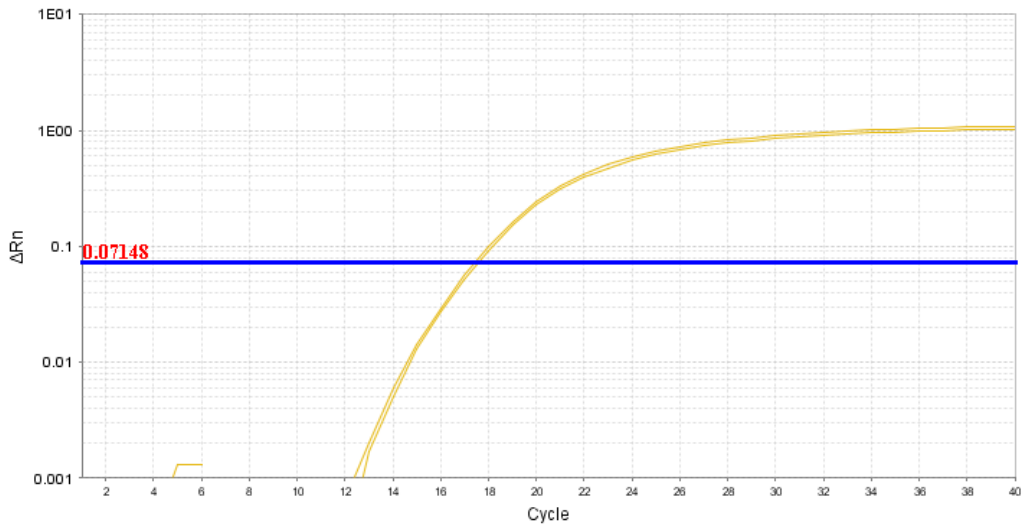


Figure 9. SN probe with 6404 (*sosorum*), AHP-3079 (*nana*)

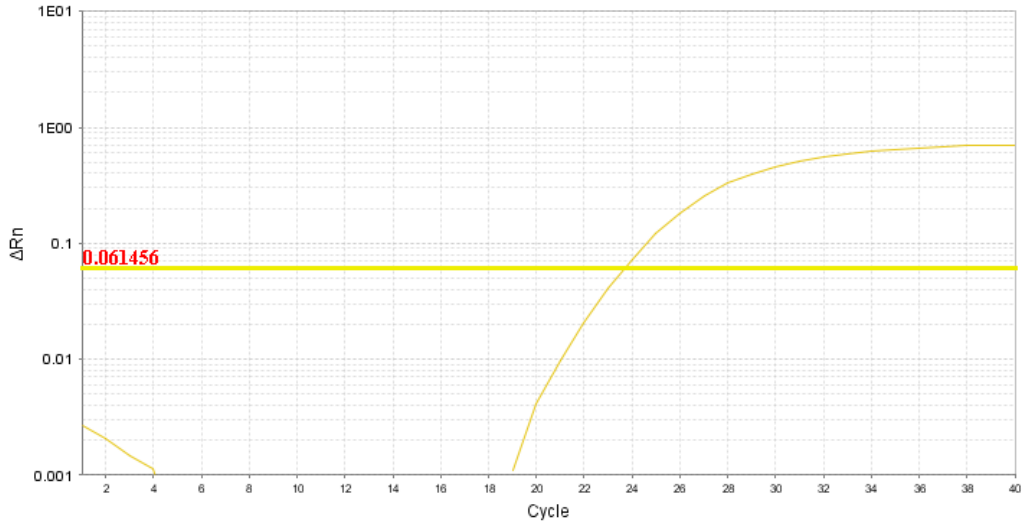


Figure 10. S1 probe with DMH-9151 (*sp. 1*)

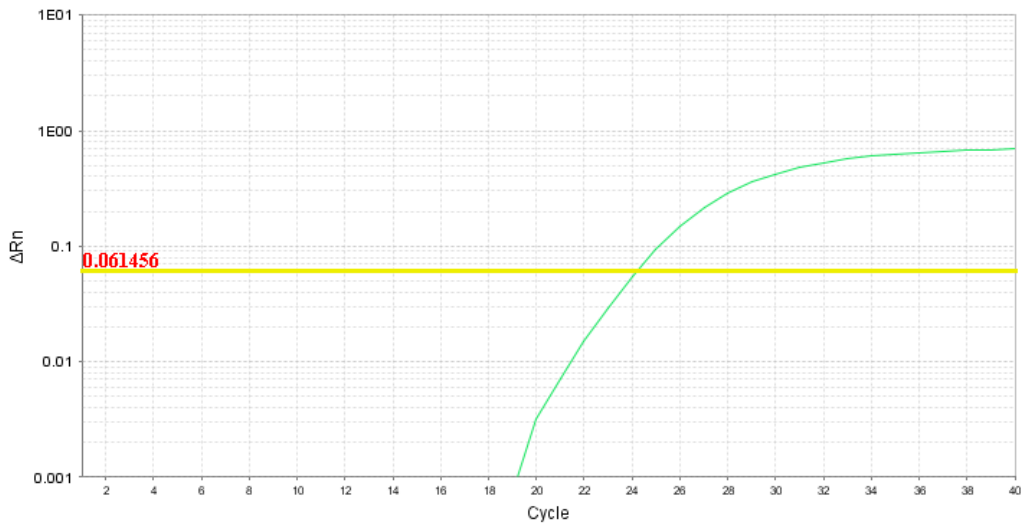


Figure 11. Probe S2 with AHP-3023 (*pricei*)

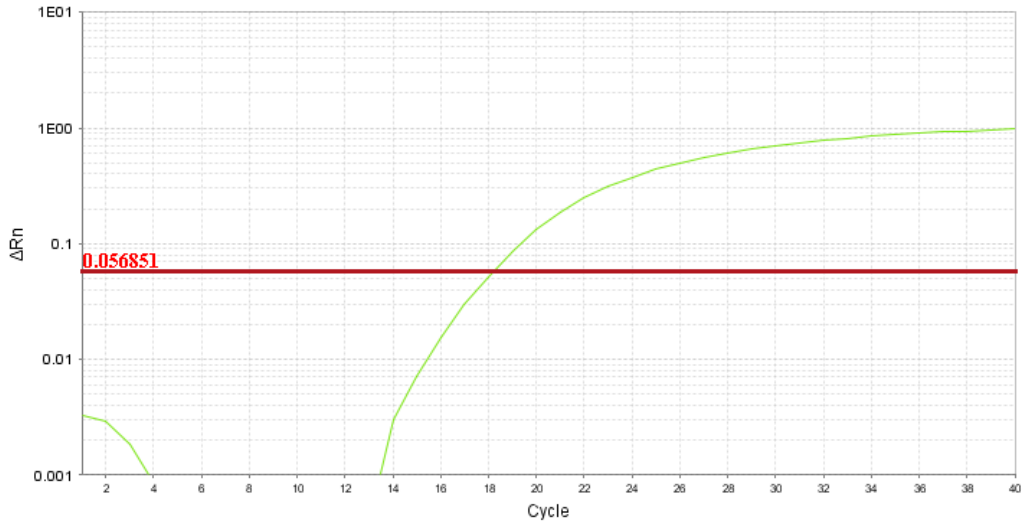


Figure 12. Probe S3 with JKK-20041 (*sp. 3*)

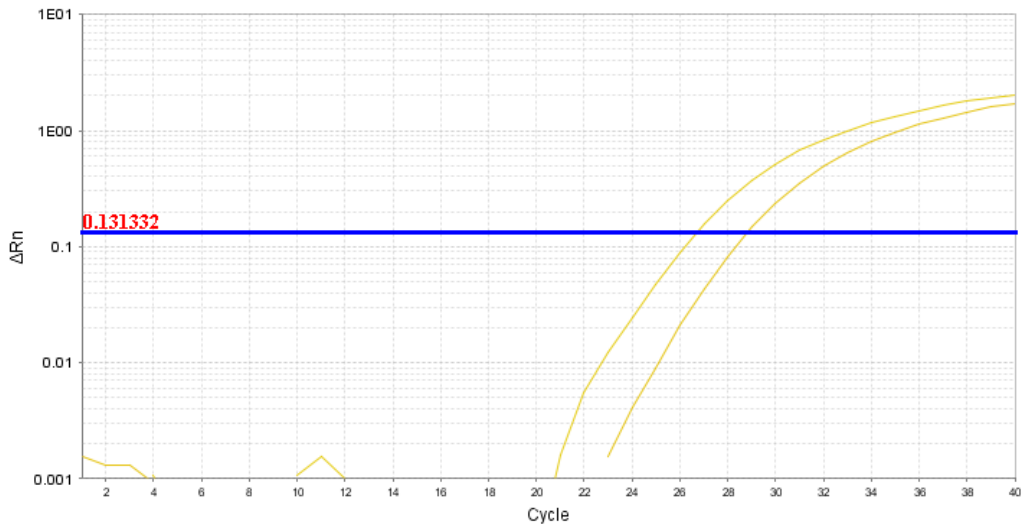


Figure 13. Probe TC with AGG-1850 (*tonkawae*), AGG-1885 (*chisholmensis*)

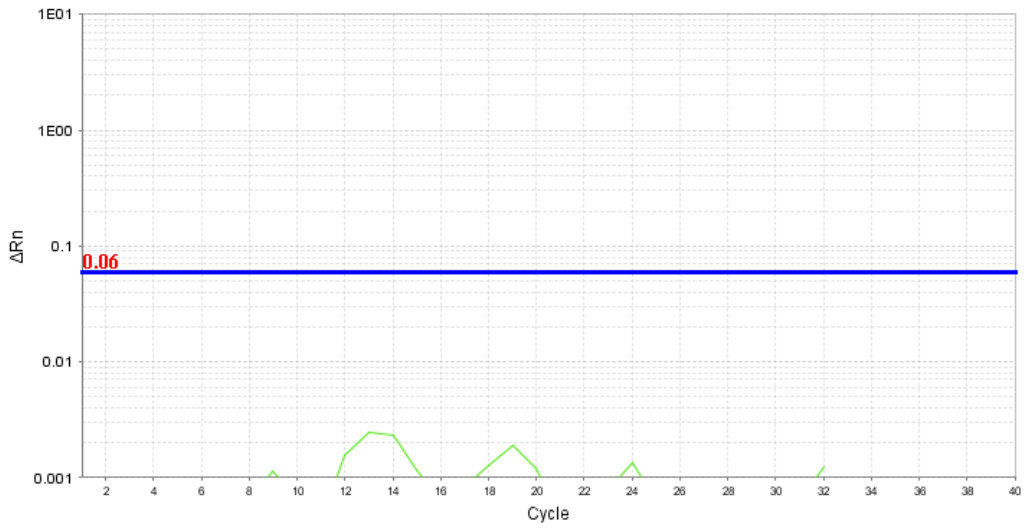


Figure 14. Probe TC with AGG-2020 (*naufragia*), Note amplification below threshold

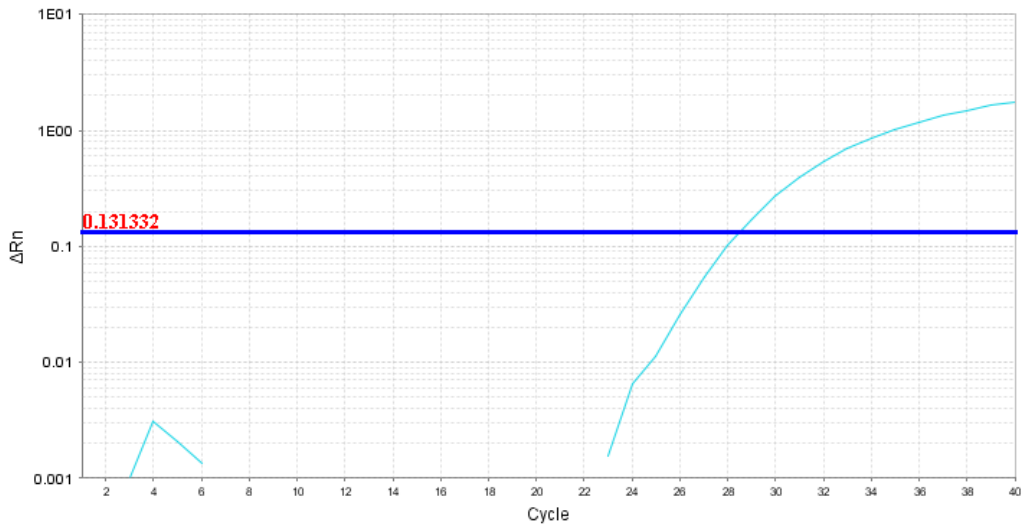


Figure 15. Probe T with TNHC-60312 (*troglodytes*)

Environmental Assays

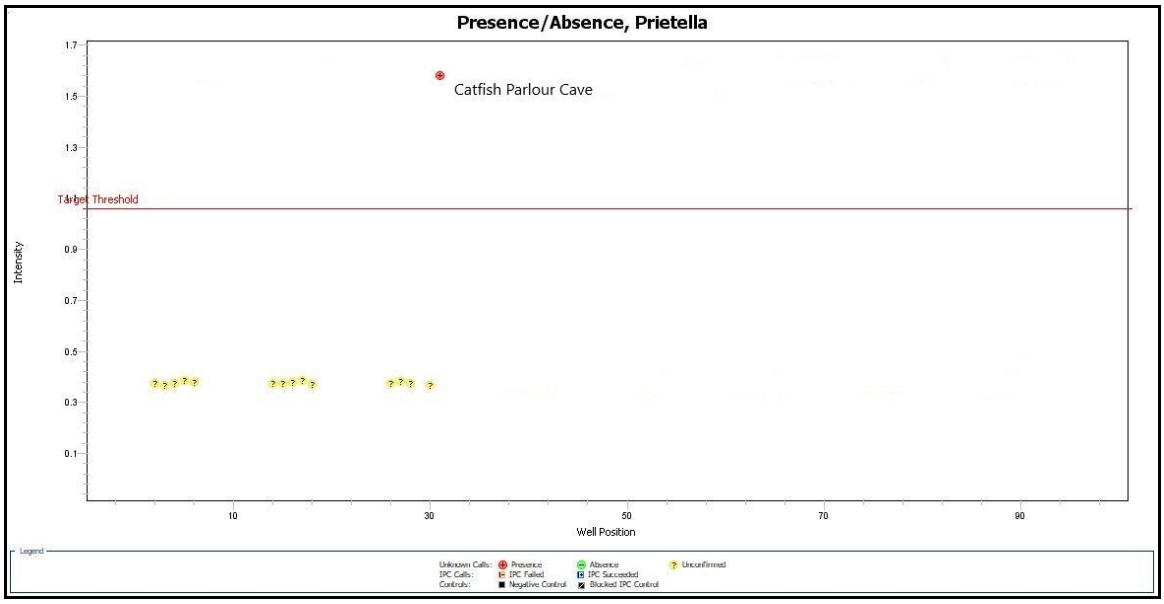


Figure 16. *Prietella* assay, positive amplification marked

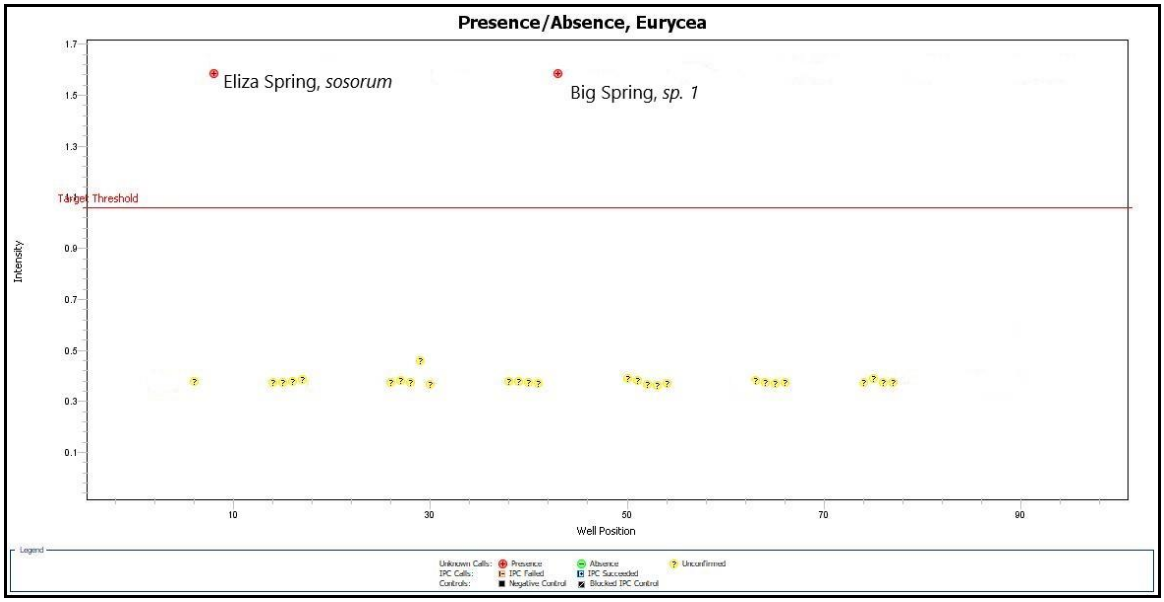


Figure 17. *Eurycea* assay, two positive amplifications marked

References

Bendik, N.F., Meik, J.M., Gluesenkamp, A.G., Roelke, C.E., and P.T. Chippindale. 2013. Biogeography, phylogeny, and morphological evolution of central Texas cave and spring salamanders. BMC Evolutionary Biology 13:201.

Devitt, T.J., Wright, A.M., Cannatella, D.C., and D.M. Hillis. 2019. Species delimitation in endangered groundwater salamanders: implications for aquifer management and biodiversity conservation. Proceedings of the National Academy of Sciences USA 116:2624-2633.

Hendrickson, D.A., Krejca, J.K., and J.M. Rodríguez Martinez. 2001. Mexican blindcats genus *Prietella* (Siluriformes: Ictaluridae): an overview of recent explorations. Environmental Biology of Fishes 62: 315-337.

Hendrickson, D.A., et al. Discovery of the Mexican Blindcat, *Prietella phreatophila*, in the U.S., and an update on its rangewide conservation status. Meeting of the Texas Academy of Sciences, 2017.

Lindgren, R.J., et al. Conceptualization and simulation of the Edwards Aquifer, San Antonio region, Texas. USGS Scientific Investigations Report 2004-5277.

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R.M., and K.C. Gough. 2014. The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. Journal of Applied Ecology 51:1450-1459.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna Austria. <https://www.R-Project.org>

Renshaw, M.A., Olds, B.P., Jerde, C.J., McVeigh, M.M., and D.M. Lodge. 2015. The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction. Molecular Ecology Resources 15:168-176.

Thomsen, P.F., and E. Willerslev. 2015. Environmental DNA - an emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4-18.

Turner, C.R., et al. 2014. Particle size distribution and optimal capture of aqueous microbial eDNA. Methods in Ecology and Evolution 5:676-684.

Wilcox T.M., et al. 2013. Robust detection of rare species using environmental DNA: the importance of primer specificity. PLOS ONE 8:e59520.

Wright, E.S. 2016. Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. The R Journal 8(1):352-359.